Intervalle de fluctuation

Présentation du sujet :

On sait qu'une machine qui produit des pièces de précision respecte les contraintes fixées pour 60% des pièces produites, les autres pièces étant défectueuses. On souhaite déterminer le nombre de pièces qu'il faut produire pour assurer la commande d'un nombre donnée de pièces dans 95% des cas.

Exposé des motifs :

Les élèves doivent rencontrer en classe de seconde la notion de fluctuation d'échantillonnage. Ils ont vu par exemple que lorsque l'on lance 200 fois un dé équilibré, la fréquence des 6 fluctue entre deux séries de lancers.
En lançant 200 fois un dé pipé, la fréquence des 6 observée ne sera pas égale à la probabilité de faire 6 avec ce dé, car en procédant à d'autres séries de 200 lancers, il y aurait fluctuation d'échantillonnage. Cependant, la fluctuation d'échantillonnage obéit à des lois : lorsque l'on considère un échantillon de n personnes tirées au hasard dans une population qui comporte une proportion p de personnes ayant le caractère étudié, il y a une probabilité de 0,95 que la fréquence observée du caractère dans l'échantillon soit dans l'intervalle $\left[p-1,96\times\sqrt{\dfrac{p(1-p)}{n}};p+1,96\times\sqrt{\dfrac{p(1-p)}{n}}\right]$.
car on peut modéliser le fait d'interroger une personne par une épreuve de Bernouilli $b(p)$*, et (si la population est suffisament importante) le fait d'interroger un échantillon de n personnes par une loi binomiale* $\mathrm{B}(n,p)$*. Si l'on étudie la variable aléatoire* (F) *qui prend la valeur de la proportion du caractère dans l'échantillon, on obtient que son espérance est p et que sa variance est* $\dfrac{p(1-p)}{n}$*. Le théorème de la limite centrée permet d'affirmer que lorsque la loi de* (F) *pourra être ajustée par* $\mathcal{N}\left(p;\sqrt{\dfrac{p(1-p)}{n}}\right)$*. D'où le fait que* $\mathrm{P}\left(\mathrm{F}\in\left[p-1,96\times\sqrt{\dfrac{p(1-p)}{n}};p+1,96\times\sqrt{\dfrac{p(1-p)}{n}}\right]\right)=0,95$.

Ainsi pour n assez grand, il y a une probabilité de 0,95 que la fréquence de l'échantillon appartient à $\left[p-\dfrac{1}{\sqrt{n}};p+\dfrac{1}{\sqrt{n}}\right]$ (car $\sqrt{p(1-p)}\leqslant\dfrac{1}{2}$). L'intervalle $\left[p-\dfrac{1}{\sqrt{n}};p+\dfrac{1}{\sqrt{n}}\right]$ étant appelé intervalle de fluctuation au seuil de 95%.

Déroulement de la séquence :

Cette séquence propose de conjecturer en classe l'existence de l'intervalle de fluctuation et les bornes de cet intervalle pour le problème proposé. Le tableur sera utilisé avec videoprojecteur devant la classe.

Sur la première feuille de calcul, le professeur évoque avec sa classe comment on peut simuler le tirage d'un individu au hasard (ici des pièces de machine). Il saisit la probabilité de produire une pièce convenable en C6 et la formule permettant la simulation en C9 (=ENT(ALEA()+C6) ou =SI(ALEA()<C6 ;1 ;0)). Puis en pressant la touche F9, le professeur fait fonctionner la simulation. *Il est possible à ce stade d'évoquer le problème des références absolus que l'on rencontrera si l'on copie la formule vers le bas.*

Ensuite dans la deuxième feuille de calcul, le professeur présente explique que cette feuille est la simulation de 250 échantillons de n productions de pièce. Ce nombre n étant fixé dans la cellule Q2. (Pour cela le professeur saisi un nombre entier entre 0 et 500 (300 pour le fichier _ligth), puis il presse la touche F9 pour lancer la simulation. Le professeur fait ensuite remarquer que pour chaque échantillons de n pièce, la fréquence des pièces convenables est calculée dans la ligne 507).

Lorsque les élèves ont compris ce que fait la deuxième feuille de calcul, le professeur va utiliser la troisième feuille de calcul pour faire conjecturer l'existence de l'intervalle de fluctuation au seuil de 95%. Pour cela, il choisit une taille n d'échantillon dans la cellule Q2 de la second feuille de calcul, puis dans la troisième feuille de calcul, par tatonnement en saisissant un réel dans la cellule C5. Il détermine la valeur du rayon r de l'intervalle centré en $0,6$ qui contiendrait "la plupart du temps" 95% des fréquences des 250 échantillons simulés dans la deuxième feuille. Lorsqu'une valeur de r est conjecturée pour une taille n d'échantillon. Le professeur reporte n et r dans le tableau ligne 16 et 17. On procède ainsi pour plusieurs tailles d'échantillon (25, 50, 100 ...). Enfin, en affichant les images des valeurs de n de la ligne 16 dans la ligne 18 par la fonction $x \mapsto \frac{1}{\sqrt{x}}$ ou les images des valeurs de r de la ligne 17 dans la ligne 18 par la fonction $x \mapsto \frac{1}{x^2}$, le professeur fait conjecturer par la classe la formule donnant l'intervalle de fluctuation en fonction de la taille des échantillons.

Compétences mises en jeu :

TICE :

- Réalisation dans une feuille de calcul automatisé d'une simulation.
- Utilisation de formules ALEA(), ENT(), SOMME(), SI() et NB.SI().
- Maîtrise des références absolues et relatives.
- Obtention d'un polygone des effectifs à l'aide d'un tableur.

Mathématiques :

- Notion statistique de classe, de fréquence, d'effectif, d'effectif cumulé.
- Notion de probabilité : expérience aléatoire, intervalle de fluctuation..

Remarques concernant les documents fournis :
Le fichier de tableur est fourni en trois versions :
- *le fichier "fluctuation" qui est le fichier à utiliser en classe,*
- *le fichier "fluctuation_light" qui est le même que le précédent mais qui limite la taille de l'échantillon à 300 pièces (contre 500 avec le fichier précédent). Ce fichier est fourni pour le cas où l'ordinateur utilisé n'est pas assez puissant pour des échantillons de 500 pièces.*
- *le fichier "fluctuation_corr" qui est la version complètée du premier fichier.*

Forme du travail :

Travail de travail en classe au videoprojecteur.

Logiciels pouvant être utilisés :

Tableur.

Niveaux possibles :

Classes de seconde.